DOCUMENT RESUME

ED 208 204 CE 030 318

AUTHOR Dickinson, Gary
TITLE Introduction to Teaching Adults. Module 4: Using Evaluation Instruments.
INSTITUTION British Columbia Univ., Vancouver. Center for Continuing Education.
SPONS AGENCY British Columbia Dept. of Education, Victoria.
PUB DATE 81
NOTE 25p.; For related documents see CE 030 315-325.
AVAILABLE FROM University of British Columbia, Centre for Continuing Education, 5997 Iona Dr., Vancouver, BC, V6T 2A4 ($40.00 per set, including cassettes, for Canadian residents; $110.00, plus postage and handling, for international orders; 10% discount on 10 or more sets).

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Adult Education; *Adult Educators; Attitude Measures; Autoinstructional Aids; Behavioral Objectives; Check Lists; Continuing Education; Correspondence Study; Learning Activities; Learning Modules; Material Development; Measures (Individuals); Rating Scales; *Student Evaluation; *Teacher Education; *Test Construction; Testing; Tests

ABSTRACT

This module on using evaluation instruments is the fourth of 11 modules in the set, Introduction to Teaching Adults. Designed to meet the learning needs of part-time continuing education instructors, these modules can be used as resource materials for local workshops or study-discussion groups, as self-instruction (each module takes approximately two hours), and as a correspondence course. (Source information on correspondence study is cited in module 1.) All modules except the first are designed to be used in part in conjunction with an audiotape. The purpose of this module is to introduce some basic concepts about evaluating learning and constructing evaluation instruments. The booklet first provides some notes from the lecture on the cassette. It then presents detailed instructions and examples to aid teachers in constructing different types of evaluation instruments, including tests, checklists and rating scales, and attitude measures. (YLB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED208204

INTRODUCTION TO TEACHING ADULTS

MODULE 4

USING EVALUATION INSTRUMENTS

Prepared by:
GARY DICKINSON
Division of Adult Education
The University of British Columbia
1981

CE 030 318

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

J. Kulich

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

MODULE 4: USING EVALUATION INSTRUMENTS

Prepared by:
The Centre for Continuing Education
The Univeristy of British Columbia
5997 Iona Drive
Vancouver, B.C., Canada V6T 2A4

Under contract to:
The Continuing Education Division
Ministry of Education
835 Humboldt Street
Victoria, B.C., Canada V8V 2M4

Design, typesetting and layout:
the strait media group inc.

Copyright © 1981 by the Ministry of Education.
All rights reserved.

No part of this book may be reproduced, stored in a retrieval system, or transcribed, in any form or by any means - electronic, mechanical, photocopying, recording, or otherwise - without prior permission of the publisher.

MODULE 4

Important Note

If you intend to submit the assignment for evaluation, please include this form with your submission and mail to:

Adult Education Programs,
Centre for Continuing Education,
5997 Iona Drive,
University of British Columbia,
V6T 2A4

Name _____

Address _____

Telephone Business _____ Home _____

Date Submitted _____

ABOUT THIS PROGRAM

This series of modules was developed by the Division of Adult Education and the Centre for Continuing Education of the University of British Columbia under contract with the Ministry of Education of the Province of British Columbia. The series was developed under the direction of Gary Dickinson, and Sandy McGechaen supervised the production.

The eleven modules in the complete program are:

Level I

1 Guidelines for Teachers of Adults

Level II

2 Using Instructional Techniques
3 Using Instructional Media
4 Using Evaluation Instruments

Level III

5 Designing Instruction
6 Writing Instructional Objectives
7 Managing Instruction
8 Enhancing Motivation

Level IV

9 Understanding the Adult as Learner
10 Understanding the Learning Process
11 Continuing Your Development

MODULE 4: USING EVALUATION INSTRUMENTS

CONTENTS

Page

Introduction 1
Notes from the Cassette 2
Procedure for Developing Instruments 3
Tests 5
Checklists and Rating Scales 11
Attitude Measurement 17
Assignment 19

INTRODUCTION

The purpose of this module is to introduce some basic concepts about evaluating learning and constructing evaluation instruments. When you have completed all of the activities in the module, including the assignment, you should be able to construct evaluation instruments to measure learning outcomes in the courses you teach.

The cassette for this module is a lecture about basic concepts and principles in evaluating learning. It runs for about 25 minutes, and you should listen to the cassette before you read the booklet. The booklet provides some notes from the lecture, so you need not bother with note-taking while listening to the tape. The booklet also includes detailed instructions and examples to aid you in constructing different types of evaluation instruments.

Begin by listening to the cassette

NOTES FROM THE CASSETTE

Evaluation — the process of judging the value or worth of something.

Formative Evaluation — evaluation conducted while a course is in progress to enable difficulties to be detected and corrected.

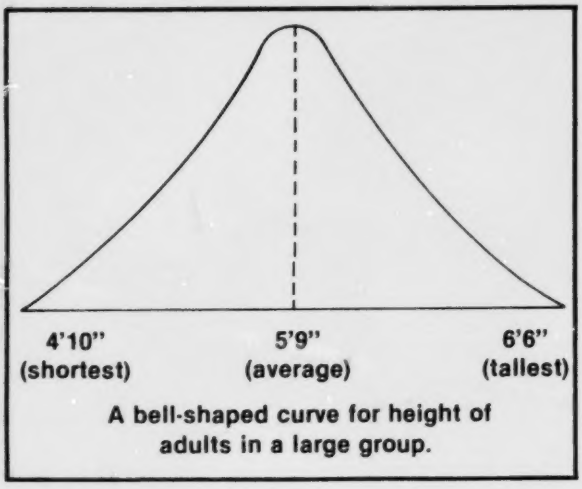
Summative Evaluation — evaluation conducted at the conclusion of a program to "sum up" the learning that occurred.

Uses of Formative Evaluation:

1. To provide feedback to teachers
2. To provide quality control
3. To predict summative evaluation results

Uses of Summative Evaluation:

1. To assign grades or certify skills and abilities
2. To predict success in future courses
3. To determine a starting point in future courses
4. To compare results in different groups
5. To provide feedback to students.



A bell-shaped curve for height of adults in a large group.

Elements in an Evaluation:

1. Criteria — standards of performance as stated in the objective
2. Evidence — the information obtained about learner performance.
3. Judgment — a decision about whether performance is satisfactory

Norm-referenced Evaluation — an approach to evaluation which compares an individual against everyone else in the relevant group.

Criterion-referenced Evaluation — an approach to evaluation which considers individual performance in relation to a criterion or standard.

Characteristics of Good Evaluation Instruments:

1. Content Validity — measurement of the specific behavior stated in the objective, not some other behavior
2. Predictive Validity — the relationship between performance on an evaluation instrument and performance in "real-life"
3. Reliability — the extent to which an evaluation instrument gives consistent results
4. Objectivity — scores obtained should not be influenced by the person who marks the test
5. Practicality — an instrument should not be too costly in terms of time or money to develop and administer
6. Interest — evaluation instruments should be interesting to the students who complete them

PROCEDURE FOR DEVELOPING INSTRUMENTS

The process of evaluating learning may seem to be complex and difficult at this point, but it can be made much easier if you approach it systematically. Following is a step-by-step procedure that should help to simplify the process.

1. Develop a General Evaluation Plan:

A plan for evaluating learning provides an overview of how evaluation will fit into the management of your course. Before starting to teach, you should decide such questions as:

- How much time should be spent on evaluation activities during the course?
- Should participants be pre-tested when they enter the course?
- What kinds of formative and summative evaluation are necessary?

Decisions about those matters should be written down to form your general plan for evaluation.

2. Review the Course Objective:

When you have decided how evaluation fits into your course, review the instructional objectives you have prepared. Make sure that the objectives are in behavioral terms so that it will be possible to measure their attainment. All of the objectives are not equal in importance, so assign weightings to them that reflect their priority. For example, if objective No. 5 is the least important, give a weight of 1. If objective No. 3 is twice as important as objective No. 5, assign it a weight of 2, and so on. These weightings will be helpful later when you construct your instruments because they indicate areas of emphasis. In general, you'll want to do a more thorough job of evaluating the most important objectives to ensure they are mastered.

3. Determine the Instruments Needed:

Based on the review of your objectives, you should be able to specify the instruments needed to conduct the evaluation. For learning outcomes in the motor skills area, you will probably use checklists or rating scales. Where attitudes are the learning outcomes, you could decide to construct attitude scales or to observe the learners' behavior. Verbal information, intellectual skills, and cognitive strategies are usually measured by tests, which may be selected responses, constructed responses, or a combination of the two types. **Selected response** means that the learner chooses the correct response from several which are presented, such as true-false or multiple-choice items. **Constructed response** items require the participant to create an answer, as is the case for essay and one-sentence answer items, or sentence completion.

4. Prepare the Items for the Instruments:

When you have reviewed the objectives and identified the instruments needed, begin writing test items or scales in draft form. Take each objective in turn and prepare the items required to assess learner attainment of the objective. Detailed guidelines on how to write items are presented later in this booklet.

5. Edit, Assemble, and Reproduce the Instruments:

With your first draft completed, edit the items for grammar, sentence structure, vocabulary level, and style. Try to place yourself in the learners' position and see if you could complete the items correctly if you were completing the form. Arrange the items in a logical order. The first items in a test should be easier than later items, to build up the learners' confidence. Try to have a mixture of different types of items so that the test isn't boring. Make sure that you give adequate directions about the kinds of response the participants must make. When you're satisfied that the instrument is in its best form, arrange to have it reproduced in sufficient quantity. The administrator of your course should be able to arrange reproduction of tests and scales, but make sure you allow sufficient time for this.

6. Administer the Instrument:

As you proceed through your course, evaluation instruments will be administered at various times as indicated in your general plan. In administering them, keep in mind the reduction in the speed capabilities of older learners so that you allow enough time for instruments to be completed. Remember also to try to reduce the threat and anxiety that evaluation can arouse in adult learners. Point out the purpose of the instrument and describe how it fits into your overall instructional plan. Tell the participants in advance the types of items that will be on a test, and how they can prepare for it.

7. Analyze the Results:

When an instrument has been administered and scored, there are several ways of analyzing the results. One is to review the performance by each learner on the items relating to each objective. This gives you detailed information about the extent that each person has attained each objective, and will identify points requiring additional instruction. There are some simple statistics which can be calculated to describe the performance of a group of students, including the range, the median, and the mean or average. Let's assume, for example, that 15 students completed a 20-item test, and the scores were: 20, 19, 19, 17, 17, 17, 16, 15, 14, 13, 12, 8, 7, 5, 4. The **range** would be 20 — 4 = 16, which is the difference between the highest and lowest scores. The **median** or middle score would be 15. The **mean** or average score is the total of the scores divided by the number of students, in this case = 13.5. At this point you could also examine your instrument in terms of the characteristics of good instruments as described on the cassette. For example, you could investigate reliability by comparing the scores on odd-numbered items with scores on even-numbered items.

8. Revise and File the Instruments:

The results obtained from analyzing the instruments should be used to revise and improve them. Errors caused by the wording of some items could be improved. Some items could be made easier and others harder to answer. The revised items can be written on a 3" x 5" or 5" x 8" index card and filed for future use. When you teach the course again, you can draw upon your file of accumulated test items for new instruments.

TESTS

Many different types of items are used to test verbal information, intellectual skills, and cognitive strategies. The two basic types are selected response and constructed response items, with the difference being the way that the learner handles each type. **Selected response** items require the learner to choose one answer from several provided, and this is usually a test of recognition memory. Instructional objectives with words like "select", "recognize", "identify", and "match" are indications that selected response items are the best to use. **Constructed response** items require the learner to create an answer, and this is usually a test of recall memory. Objectives including behaviors such as "list", "construct", "compare", and "evaluate" are indications that constructed response items should be used.

In comparing the two types of items, selected response items are:

1. More difficult for the instructor to write
2. Easier to score
3. More objectively scored
4. More subject to guessing by the participant
5. Quicker to complete and,
6. More extensive in their coverage.

Constructed response items, however, are:

1. Easier for the instructor to write
2. More difficult to score
3. More subjective in their scoring
4. Less subject to guessing by the participant
5. Longer to complete and,
6. Limited in their coverage of a subject.

Selected Response Items

Three commonly used kinds of selected response items are multiple-choice, true-false, and matching items. Following are some guidelines and examples for each type.

A. Multiple-choice:

The multiple choice item is perhaps the best known type of selected response. These items present more than one alternative answer for the student to select. There are several rules to follow when writing multiple-choice items.

1. Present the problem in a clear and unambiguous form. Compare these two examples:

Poor	**Good**
Plants live because of:	The biochemical process by which plants sustain life is called:

The second item is much less ambiguous than the first one. Without even looking at the alternatives, the student can tell immediately what kind of answer will be correct.

2. Avoid the use of specific determiners. A specific determiner is characteristic of a poorly written item that tends to give away the answer. Compare the following examples:

Poor	**Good**
The purpose for wearing heavy gauntlets when welding is to:	The purpose for wearing heavy gauntlets when welding is to:
a. Prevent electric shocks	a. Prevent painful electric shocks
b. Keep your hands cool	b. Prevent loss of grip
· c. Prevent your hands or wrists from getting burned from sparks or flying metal particles.	· c. Prevent painful burns

The first example is made artificially easy because the correct response is so much longer than the others. In the same way, if the correct answer is very short when compared to other alternatives, it may be a giveaway.

Poor	**Good**
The proper utensil to use for baking biscuits is a:	The proper utensil to use for cooking biscuits is a:
· a. Baking sheet	· a. Baking sheet
b. Dutch oven	b. Dutch oven
c. Double boiler	c. Double boiler
d. Cooking pan	d. Cooking pan

In the first example, the word "baking" appears in both the item stem and in the correct response. The uninformed student could successfully guess the correct answer on the basis of this similarity. By changing the word 'baking' to 'cooking', the instructor can avoid this problem. The uninformed student following the same strategy for this item would choose the wrong answer, 'cooking pan'.

Poor

The test instrument used to measure electromotive force is a:

a. Ohmmeter
•b. Voltmeter
c. Ammeter

Good

The test instrument used to measure electromotive force is:

a. An ohmmeter
•b. A voltmeter
c. An ammeter

In the first example, the correct response is the only response that grammatically matches this flaw with a simple change.

3. Avoid the use of "hang-on" alternatives. There are very few occasions in which alternatives such as "all of these" and "none of these" are required. Compare the following examples:

Poor

What instrument(s) can be used to find the mid-point of a 9-inch steel bar?

•a. Calipers
b. A straightedge
c. A protractor
d. All of the above
e. None of the above
f. a and c only

Good

What instrument(s) can be used to find the mid-point of a 9-inch steel bar?

•a. A caliper
b. A straightedge
c. A protractor

There is nothing sacred about maintaining any given number of response alternatives for all multiple-choice items. An item-writer should not struggle to produce, say, four alternatives for each item. The first example above wastes the student's time by requiring more reading than is necessary. More often than not, the use of "all of the above" or "none of the above" is not required.

4. Avoid redundant reading in the items. Often, the amount of reading required by the student, as well as the space taken up by an item can be reduced. Compare these two examples:

Poor

Good corn seed selection can be accomplished by:

a. Shucking and planting those seeds from the least populated stalks
b. Shucking and planting those seeds from the tallest stalks
•c. Shucking and planting those seeds from the most populated stalks.
d. Shucking and planting those seeds from the shortest stalks

Good

Good corn seed selection can be accomplished by shucking and planting those seeds from the:

a. Least populated stalks
b. Tallest stalks
•c. Most populated stalks

Both examples pose the same problem, but the second item is much shorter and easier to read.

5. Keep numerical alternatives in a logical order. Items which have only numerical alternatives are easier to read and respond to if the alternatives are in a logical order. Compare the following examples:

Poor	Good
What is the specific gravity for a battery with a hydrometer reading of 1.220 at 90° F?	What is the specific gravity for a battery with a hydrometer reading of 1.220 at 90° F?
a. 1.220	a. 1.216
b. 1.320	b. 1.220
c. 1.260	*c. 1.224
d. 1.216	d. 1.260
*e. 1.224	e. 1.320

The second item required less reading for the student. Notice that the item alternatives could also have been placed in descending order without confusing the student.

B. True-False:

True-false items are special types of multiple-choice items because they offer only two alternatives. These items can usually be answered even more quickly than an ordinary muitiple-choice item. There are several rules to follow for constructing true-false items:

1. Be sure the item is absolutely true or false. The student who is aware of exceptions to a particular statement will be confused as to which way to respond to the item.

2. Avoid the use of specific determiners. Many students are aware that a statement is likely to be false if it contains the words 'always' or 'never'. To include such words serves little purpose.

3. Avoid the use of negatives or double negatives. These tend to make the item more difficult than it would be otherwise. Compare the following examples:

Poor	Good
It is not correct to say that the dodo was unable to survive past the 17th century.	The dodo became extinct during the 17th century.

A half-dozen items like the second one could be answered in the time that the student would use to decipher the first item.

4. Include a larger number of false items than true ones. Students who guess on true-false items tend to select true more often than they select false. This fact unfairly favours the uninformed student who is given a test primarily composed of true items.

C. Matching Items:

A matching item is actually a condensed set of multiple-choice items, each having the same set of response alternatives to choose from. There are two basic rules to follow in constructing matching items:

1. Make the directions for the item clear and complete. Compare these two examples:

Poor	Good
Match the following:	Match each answer listed on the right with the proper statement on the left. Some answers may be used more than once.

The second set of directions more clearly sets the task for the student.

2. Use more alternatives than there are items to match. Or, allow answers to be used more than once. This prevents the student from obtaining correct answers by the process of elimination.

D. General Guidelines:

There are some general guidelines that would be useful to keep in mind when writing any type of selected response test item.

1. Provide clear directions so the student knows how to respond to the item.

2. Keep the vocabulary level of the items as simple as possible. Unless the test is meant to measure reading comprehension, there is no need to use three syllable words when shorter ones will suffice.

3. Write the items so that they present the task or pose the problem to the student in a clear and unambiguous manner.

4. Keep the items independent of each other. Avoid writing a series of items in which the answer to one item determines how the following item will be answered.

5. Write items covering important concepts, not trivial information. The surest way to construct a test measuring the acquisition of trivial information is to write the items without allowing time for careful planning.

6. Avoid trick questions. Unless you are trying to teach students how to cope with trick questions, including them can only serve to confuse the students.

7. Avoid writing items which give the answer to other items in the same test. Such items obviously interfere with your measurement of what the student knows.

8. The answers to selected response items should be those on which other content experts can agree. If no consensus can be obtained, the item should be rewritten.

Constructed Response Items:

Two commonly used kinds of constructed response items are those involving short answers and essays.

A. Short Answer:

Short answer items require a response ranging from a few words to a complete sentence. The completion question is a type of short answer item that calls for one or more responses to be inserted in order to complete a sentence or phrase. There are several rules to follow in the preparation of short answer items.

1. Include directions for the student to follow. Should the student prepare a complete sentence, or will a few key words suffice? Should the student write in the blanks in a question or underneath the item? These are simple considerations, but they are important because they let the student know how to respond.

2. State the item as precisely as possible. This is an especially critical point for completion items. Compare the following examples:

Poor	Good
Plants live because of:	The biochemical process by which plants sustain life is called:

A student might feel justified giving an answer to the first item such as fertilizer, water, and sunlight. The second example is less likely to be misinterpreted.

Poor	Good
What does a pilot tube work with?	With what source is the pilot tube intake compared for operating certain aircraft instruments accurately?

The first example is ambiguous. The second example more clearly defines the task for the student. The purpose of well-written test items is to determine what the student knows, not how well the student can decipher confusing test items.

Try to avoid writing completion items requiring multiple responses. In general, such items are confusing to the student, as well as being far more difficult to write well.

3. Prepare answer keys for the items. When preparing answer keys for short-answer items, try to include all the possible responses that could be considered correct. This will prevent a lengthy scoring period, since you will not have to mull over new responses, trying to decide whether they are acceptable.

B. Essay:

Essay items require a response ranging from a few sentences for short essay items to several paragraphs for extended essay items.

1. Include directions for the student to follow. Some instructors prefer that essay answers be written in complete sentences, while others prefer outline form. Be sure to make the directions explicit, so the student will know what type and length of response is called for.

2. Include guidelines for responding and scoring. It is often helpful to explain how a student's response will be evaluated. Will spelling count? Will points be added or subtracted for overall appearance? How many specific points or examples are to be included in the answer? Compare these examples:

Poor	Good
Explain bookkeeping. Use both sides of this sheet if necessary.	In 500 words or less, compare and contrast the single-entry and double-entry methods of bookkeeping. List at least two advantages and two disadvantages of

each. Spelling and neatness will not affect your score. Be sure to write complete sentences. A complete answer is worth 10 points. 2½ points will be deducted for each missing advantage or disadvantage.

The first example is so incomplete that any student would be hard-pressed to produce an acceptable response. The second example defines the task much more completely.

3. Prepare a model answer sheet for use in scoring. Answer the item yourself, being sure to include all the specific facts and examples that make up an acceptable response. Then, compare your answers to the question. They should match in terms of length, number of examples, and other features. Also, having the model answer is helpful for actually rating student papers. By making the testing and scoring process more objective, some of the hidden sources of bias associated with essay items can be reduced.

C. General Guidelines:

Several guidelines should be kept in mind when developing constructed response items.

1. Always provide clear directions explaining how to respond to the items.

2. Keep the vocabulary level of the items as simple as possible.
3. Write the items so that they present the task or pose the problem in a clear and unambiguous manner.

4. Prepare sample answers to check the clarity of the items and to use as a model for grading students' papers.

CHECKLISTS AND RATING SCALES

Checklists and rating scales are used to evaluate motor skill learning. Both types of instrument are observational tools used by the instructor to record the presence or absence of behaviors or characteristics. Some checklists and rating scales focus upon the **process** of an activity, which means that the items on the instrument would be listed in the sequence in which the action is performed. For example, a checklist could be developed to evaluate the process of parking a car or making a cake. Other checklists and rating scales focus on the **product** resulting from an activity, so the items on the instrument indicate attributes or characteristics that would be inspected. Thus, the evaluation of the cake would be in the eating rather than in observing how it came to be produced. Some product rating scales provide several examples of products of varying quality, and the learner's product is compared with different specimens to determine which one it closely resembles.

Checklists

Checklists are forms for recording the presence or absence of behaviors or characteristics. In preparing a checklist, the initial step is to identify and list the behaviors or characteristics in a logical order, usually the order in which they are performed. The items in the list should include all of the important components of the task; they should be clearly stated, and they should include action words. Here is an example of a checklist.

CHECKING A CAR ENGINE'S OIL LEVEL

Name: _____ Date: _____

Observer: _____

DID	DID NOT	ACTIVITIES
_____	_____	1. Turn car engine off.
_____	_____	2. Put car in park position.
_____	_____	3. Open hood of car and remove dipstick.
_____	_____	4. Wipe oil off dipstick.
_____	_____	5. Reinsert dipstick in hole.
_____	_____	6. Remove dipstick.
_____	_____	7. Check the oil marking.
_____	_____	8. Replace dipstick if oil level is at "add oil" mark or above.
_____	_____	9. Add oil if oil level is below "add oil" mark.

Amount of time taken: _____ minutes

Notes: _____

Several variations can be used to make a checklist more useful as an instructional tool. A numbering system can indicate a sequence of steps. It is sometimes desirable to note the amount of time taken for an action to occur. You might wish to leave some blank space to note errors, omissions, or comments about the performance. Finally, you could use a section on the form to suggest to the student some steps that could be taken to improve the performance.

Rating Scales

Rating scales are also observational tools. While a checklist just records the presence or absence of an action or characteristic, a rating scale provides a form to record the quality or level of performance. The form focusses the observer's attention on specific aspects of the performance.

The construction of rating scales again begins with the instructional objective. The important factors in successful performances are identified and listed, and descriptive or numerical scales are chosen. A numerical scale provides little more than numbers which are circled by the observer, as shown in the following example.

EXECUTION OF TENNIS SHOTS

Name: _____ Date: _____

Observer: _____

Forehand drive:
Poor 1 2 3 4 5 Excellent

Forehand lob:
Poor 1 2 3 4 5 Excellent

Backhand drive:
Poor 1 2 3 4 5 Excellent

Backhand lob:
Poor 1 2 3 4 5 Excellent

Comments: _____

Descriptive scales provide much information about the standards required for successful performance. In teaching a typing class, for example, the following descriptive rating scale might be used for rating neatness.

NEATNESS

Poor ——— Average ——— Excellent

Poor many corrected or uncorrected typographical errors, may have a cover page, margins are one inch or less, and a folder may or may not be used.

Average some corrected typographical errors, a cover page, over one-inch margins on both sides of the paper but not the margins suggested, and placed in a manila folder.

Excellent no typographical errors, a cover page, 1½ inch margins on both sides of the page, and placed in a manila folder.

In developing rating scales, you may wish to weigh some aspects more heavily than others. If you felt, for example, that being able to hit a forehand drive was three times as important as a forehand lob in the tennis scale shown earlier, then the score for the more important stroke would be multiplied by three.

Three errors are commonly made in using rating scales. Avoid rating everyone close to the average as people will tend to be spread over a wider range. Some instructors tend to over- or under-rate everyone compared with what other instructors would do. Finally, avoid being influenced by your general impression of the student and concentrate on the specific behavior under observation.

ATTITUDE MEASUREMENT

The exact measurement of attitudinal learning outcomes is more difficult than the types of learning outcomes already discussed. You can usually be satisfied if you know whether the change is positive or negative. In other words, is the learner developing a more favourable attitude toward something or is the attitude becoming less favourable? There are two basic ways of looking for evidence about attitude change. One is to ask the learner directly about the attitude, and the other is to observe behavior that might indicate what attitude exists.

Observation

Because of the difficulty of measuring attitudes directly, we sometimes choose instead to observe behavior related to an attitude and infer the attitude from that. For example, in teaching defensive driving we infer that an attitude of defensiveness exists if the student acts in certain ways. In other instances, the attitude change is not as important as we might think and the behavior is actually the crucial objective. For example, in teaching dental hygiene we are not as concerned that students develop favourable attitudes about brushing their teeth twice a day as we are that they actually do so.

There are many kinds of learner behavior that can be observed for the purpose of finding out whether attitudes are changing in the desired direction. Some examples are given here:

- whether students discuss the topic at coffee breaks
- absence or dropping out of the course
- number and kinds of questions asked in class
- arriving early or leaving late from class
- willingness to take further courses in the subject
- doing extra reading or assignments

Attitude Scales

Attitude scales are more formal ways of obtaining information about learning outcomes of this type. It is difficult to construct reliable and valid attitude scales, so it is best just to try to determine general tendencies. Although there are many different kinds of attitude scales, only two are discussed here: Likert scales and semantic differential scales.

A. Likert Scales

Likert scales (named after their inventor) are in the form of a statement followed by a forced-choice response usually on a scale of five to nine points. Two examples follow:

I find bookkeeping to be:

1. Very interesting
2. Somewhat interesting
3. Neither interesting nor uninteresting
4. Somewhat uninteresting
5. Very uninteresting

Bookkeeping is essential for the health of our economy.

1	2	3	4	5	6	7	8	9
Strongly disagree				Neutral				Strongly agree

Several items about different aspects of the attitude being measured are usually placed together to form an attitude scale. There are several other examples of Likert scales in Module 11 — Continuing Your Development.

B. Semantic Differential Scales

These scales try to detect differences in the meanings which words have for different people. Usually a concept is presented and several pairs of adjectives are listed to which the respondent reacts. For example:

As a trade, welding is:

pleasant	7	6	5	4	3	2	1	unpleasant
dirty	1	2	3	4	5	6	7	clean
valuable	7	6	5	4	3	2	1	worthless
bad	1	2	3	4	5	6	7	good

Adding up the scores for a particular concept will give an indication of the overall attitude about the object measured.

ASSIGNMENT

The basic concepts and procedures for evaluating learning discussed in this module may seem somewhat overwhelming at first, but remember that evaluation is an ongoing process. The procedure for developing instruments presented earlier in this booklet takes several years to complete, so don't be discouraged if you can't evaluate as well as you want the first time you teach a course.

While you are working on this module, it would be useful to take some time to practice construction of the various types of items and instruments presented. Based on typical instructional objectives for your course, see if you can write two or three items of each asterisked type listed below. Keep in mind the guidelines presented for each type. Submit the items you prepare in written form for comment.

TESTS

Selected response items

*Multiple-choice
*True-false
*Matching

Constructed response

*Short answer
*Essay

CHECKLISTS AND RATING SCALES

*Checklists
*Rating scales

ATTITUDE MEASURES

Observation
Attitude scales

*Likert
*Semantic differential